Bi-level heterogeneity modeling of functional performance degradation for the aging population

Nazmus Sakib, Xuxue Sun, Nan Kong, Hongdao Meng & Mingyang Li

To cite this article: Nazmus Sakib, Xuxue Sun, Nan Kong, Hongdao Meng & Mingyang Li (2017): Bi-level heterogeneity modeling of functional performance degradation for the aging population, IISE Transactions on Healthcare Systems Engineering, DOI: 10.1080/24725579.2017.1339147

To link to this article: http://dx.doi.org/10.1080/24725579.2017.1339147
Bi-level heterogeneity modeling of functional performance degradation for the aging population

Nazmus Sakib, Xuxue Sun, Nan Kong, Hongdao Meng, and Mingyang Li

ABSTRACT
As the majority of service recipients of the U.S. long-term care system, the elderly with disabilities experience heterogeneous functional limitation and performance degradation. Successful modeling of heterogeneous degradation becomes of great importance for long-term care decision making. Existing heterogeneity modeling approaches are mainly restricted to single-level modeling of functional performance degradation for older adults. This article proposes a novel bi-level heterogeneity modeling and quantification framework to systematically investigate heterogeneous performance degradation at both sub-population and individual levels. Specifically, a Bayesian non-parametric modeling and inference framework is first proposed to characterize sub-population-level degradation heterogeneity. It relaxes the conventional assumption of pre-specifying a fixed number of sub-populations and allows joint model estimation and selection. Functional data analysis techniques are further considered to quantify individual heterogeneity nested within sub-populations. Based on the extracted temporal variation signatures of degradation profiles, post-analysis is performed to identify relevant risk/protective factors in explaining individual heterogeneity. To illustrate the proposed work and demonstrate its validity, the heterogeneous cognitive degradation quantification of a sample of community-dwelling older adults from the Health and Retirement Study is provided.

1. Introduction
Experiencing functional (e.g., physiological or psychological) limitation and performance degradation, the elderly with disabilities are the majority of the service recipients of the U.S. long-term care system (LTC). National estimates suggest that the number of the elderly with disabilities in 2008–2012 was about 15.6 million, which accounted for 38.7% of the total number of U.S. older adults (He et al., 2014). As the baby boomers began turning 65 in 2011, coupled with decreased fertility and increased longevity, the U.S. will experience considerable growth in the elderly population with disabilities. To meet with the excess demand without compromising the quality of care, successful LTC preparedness and service planning become of great importance.

The LTC system encompasses a variety of settings and services to satisfy different needs of elderly people, which are mainly determined by their functional performance and disability conditions. Due to the highly varied characteristics among individuals, functional limitation and performance degradation of elderly people are heterogeneous, thus generating heterogeneous needs of LTC resources in different aspects. From the LTC network aspect, heterogeneous functional performance of elderly people will greatly affect their demands for different LTC facility settings in a LTC network, ranging from institutional settings (e.g., nursing homes (NH)), to home and community-based settings; e.g., assisted living facilities (ALF). According to national estimates (Harris-Kojetin et al., 2013), 90.9% and 86.6% of people in NH need dressing and toileting assistance due to major functional declines, compared to 44.9% and 36.8% of people requiring such assistance in ALF. In addition, 48.5% of people in NH have depression or cognitive impairment, compared to 24.8% and 39.6% of people in ALF, respectively. People with more severe functional limitations are more likely to need NH, since they can receive 24-hour skilled nursing care and a wide range of personal care assistance. In contrast, people with minor disability conditions may be more willing to choose home and community-based care settings, such as ALF, because their quality of life can be maximized by balancing their wellness, self-independence and privacy. Successful modeling and quantification of heterogeneous performance degradation over time will help develop more proactive and adaptive LTC preparedness decisions, such as better capacity planning and/or resource re-balancing strategies among multiple LTC settings in the region of interest, to meet the heterogeneously evolving needs of the aging population. Neglecting such heterogeneity may result in capacity shortage or surplus of certain LTC settings, causing negative consequences for both care providers and care receivers. For instance, due to capacity shortage of NH, people with severe physical limitations may be placed in ALF with less intensity of care. Their fall incidences and fall-related injuries will dramatically increase, resulting in an increase of health care cost and mortality rate.

From the aspect of a single LTC facility, heterogeneous functional performance will affect the utilization of different types of
nursing resources and their service times. According to relevant statistics reported by the Centers for Medicare & Medicaid Service (CMS) (STRIVE Project Study, 2005), for nursing home residents with reduced physical function, their utilization of nursing resources varies based on their total Activities of Daily Living (ADL) Score (Weiner et al., 1989) on a scale of 0–15. The score measures the resident’s independence in terms of activities in the nursing home, such as bed mobility, transfer, toilet use, eating, etc., and a larger score indicates less independence. For the residents with an ADL score of more than 14, their average nursing time per day is about 19 minutes for a Licensed Practical Nurse (LPN) and about 109 minutes for a Certified Nursing Assistant (CNA). In contrast, for residents with an ADL score of less than 2, their average nursing time per day is about 14 minutes for the LPN and about 18 minutes for the CNA. CNAs help residents with tasks such as eating, dressing, using the toilet, etc., and less independent residents have much higher utilization of CNAs. LPNs help plan and implement the residents’ care and treatment, and less independent residents have slightly higher utilization of LPNs. Successful modeling and quantification of heterogeneous performance degradation over time will help develop more targeted and proactive decision making in workforce planning and nursing resource prioritization based on the heterogeneous utilization of residents in a single LTC facility. Neglecting such heterogeneity and implementing one-size-fits-all decisions may lead to inappropriate staffing levels and worsen the quality of care for residents who need more LTC resources.

The majority of existing studies in modeling heterogeneous performance degradation of the aging population have primarily focused on single-level heterogeneity modeling; namely, the sub-population-level or the individual-level. Existing sub-population-level modeling mainly consists of subjective categorization (Lunney et al., 2003; Ferrucci et al., 1996) and group-based modeling (Liang et al., 2011; Nagin, 2005; Zimmer et al., 2012). The main objective of these methods is to categorize the performance degradation of the overall population into multiple sub-populations/groups with distinct temporal evolving patterns. Subjective categorization qualitatively categorizes sub-populations based on clinical observations, medical experience, and/or graphical visualization. Such ad hoc methods lack rigorous justification and are subject to subjective bias and/or visualization error (Nagin et al., 2010). To quantitatively categorize sub-populations, group-based modeling approaches develop finite mixture models to quantify heterogeneity in a two-step procedure. In the model estimation step, a series of mixture models with different pre-specified numbers of sub-populations are developed and estimated. In the model selection step, the most appropriate number of sub-populations is determined based on hypothesis tests such as likelihood ratio tests, or information criterion such as Akaike information criterion (AIC) and Bayesian information criterion (BIC). In summary, the aforementioned methods assume a known and fixed number of sub-populations, and either pre-determine it based on subjective knowledge or post-determine the number based on a two-step procedure. Our goal is to relax such assumptions without assuming a known number of sub-populations and learn objectively from data in a one-step procedure.

Unlike the sub-population-level heterogeneity modeling, the individual-level heterogeneity modeling approaches, such as discrete time transition models (Li et al., 2013; Wolfson et al., 2014) and growth-curve models (Lynch et al., 2002; Brown, 2010; Gill et al., 2010), characterize the individual-level evolving patterns of performance degradation with an implicit assumption of homogeneous population. In the discrete time transition modeling approaches, performance degradation can only be evaluated at pre-determined discrete time points. Growth-curve models are capable of evaluating degradation performance at any continuous time stamp and incorporate random parameters to quantify individual heterogeneity. However, they often utilize individual specific random intercept parameter and/or random slope parameter to quantify heterogeneity of initial degradation performance and/or degradation rate among individuals, respectively. The modeling accuracy may be constrained by the limited temporal evolving information extracted. In addition, there is no further investigation of how different potential influencing factors (e.g., risk and/or protective factors) will explain such individual heterogeneity. We want to investigate the underlying continuous degradation process of all individuals and extract richer temporal evolving information, such as the higher-order degradation curvature. Moreover, the potential influencing factors that may contribute to explaining such individual heterogeneity need to be identified.

To fill the research gap and systematically investigate the heterogeneous performance degradation of the aging population, this article proposes a bi-level (i.e., both sub-population-level and individual-level) heterogeneity modeling and quantification framework. Specifically, at the sub-population-level, a Bayesian non-parametric model is proposed to characterize sub-population-level heterogeneity by relaxing the assumption of pre-specifying the number of sub-populations in the conventional modeling approaches. The proposed formulation also allows simultaneously identifying the number of sub-populations and estimating sub-population specific parameters in a one-step procedure. In addition, a convenient estimation algorithm is proposed to address a series of technical challenges encountered during the model estimation. At the individual-level, a functional data analysis technique, Functional Principal Components Analysis (FPCA) (Ramsay et al., 2005), is employed to investigate the individual degradation heterogeneity within each sub-population. FPCA characterizes the individual performance degradation at a finer and continuous time scale and extracts richer temporal evolving information to improve the modeling accuracy. Furthermore, the extracted temporal evolving signatures of individual degradation profiles can be represented with concise and interpretable scores. They will facilitate the identification and quantification of different risk and/or protective factors involved in influencing the individual heterogeneity.

The remainder of the article is organized as follows. Section 2 presents the proposed systematic framework with sections 2.1–2.3 explaining details of its composing sub-modules; namely, sub-population-level heterogeneity modeling, estimation algorithm, individual heterogeneity modeling and post-analysis. Section 3 gives both a numerical case study and a real-data case study to illustrate the proposed work and demonstrates its performance. Section 4 draws conclusions.
Considering a heterogeneous population of elderly individuals, the functional performance degradation of individual \( i \) is evaluated at multiple discrete time points \( t_{ij}, j = 1, \ldots, n_i \), with performance measures \( y_{ij} \)'s. Conventional group-level heterogeneity modeling approaches (Nagin, 2005; Zimmer et al., 2012) assume that a heterogeneous population consists of a fixed number \( K \) of sub-populations with distinct evolving patterns characterized by \( m_i(t) \), where \( m_i(t) \) represents the mean functional degradation of individuals belonging to sub-population \( k \). \( y_{ij} \) can then be parametrically modeled as

\[
f(y_{ij} | \Theta) = \sum_{k=1}^{K} p_k f_k(y_{ij} | m_k(t_{ij}), \sigma_k^2),
\]

where \( f(\cdot) \) is the probability density function (pdf) of the overall heterogeneous population governed by a collection of all unknown parameters \( \Theta \). \( p_k \) is the proportion of sub-population \( k \) in the overall population and \( f_k(\cdot) \) is the corresponding pdf. \( f_k(\cdot) \) captures the sub-population-level heterogeneity and is a pdf of normal distribution; i.e., \( N(m_k(t), \sigma_k^2) \), where \( \sigma_k^2 \) quantifies the corresponding within-sub-population variability.

A major limitation of the conventional formulation in Eq. (1) is that a known and fixed number of sub-populations, \( K \), needs to be determined. It is either pre-determined based on subjective judgment and/or graphical visualization, or post-determined based on various model selection techniques, such as pairwise hypothesis testing or ranking of information criterion. Pre-determination is less objective and its accuracy is often constrained by visualization errors or subjective bias. Post-determination requires a two-step approach of estimating a series of candidate models by varying the number of sub-populations and selecting the most appropriate one. To relax such an assumption, obviate the difficulty of assuming a fixed number of sub-populations, and realize a one-step joint model of estimation and selection, a new model is formulated for characterizing the sub-population-level functional performance degradation heterogeneity and can be represented as

\[
y_{ij} | m_i(t), \sigma_i^2 \sim f_i(\cdot | m_i(t_{ij}), \sigma_i^2),
\]

\[
\forall i = 1, \ldots, N; j = 1, \ldots, n_i; (m_i(t), \sigma_i^2) \mid G \sim G, \forall i = 1, \ldots, N; G|\eta, G_0 \sim DP(\eta, G_0(\cdot)),
\]

where \( (m_i(t), \sigma_i^2) \) represents the sub-population specific mean degradation and variability associated with individual \( i \). \( (m_k(t), \sigma_k^2) \) has a probability of \( p_k \) to be selected as \( (m_i(t), \sigma_i^2) \) from a set of countable but infinite number of possible sub-populations; i.e., \( (m_k(t), \sigma_k^2)_{k=1}^{\infty} \). \( G \) is a random distribution generated from a non-parametric prior of Dirichlet process (DP), denoted as \( DP(\cdot) \). \( G_0 \) is the base distribution of DP with a positive diffusion parameter \( \eta \). To show the connection and difference between Eq. (1) and Eq. (3), a constructive representation of \( G \) (Sethuraman, 1994) is considered and represented as \( G = \sum_{k=1}^{\infty} p_k \delta_{(m_k, \sigma_k^2)} \), where \( \delta_{(m_k, \sigma_k^2)} \) is the Dirac delta measure of a point mass of \( 1 \) at \( (m_k, \sigma_k^2) \). Thus, the pdf of the overall heterogeneous population based on Eq. (3) can be
explicitly written as

\[ f(y_{ij} | \Theta) = \sum_{k=1}^{\infty} p_k f_k(y_{ij} | m_k(t_{ij}), \sigma_k^2), \]

\[ \forall i = 1, \ldots, N; \ j = 1, \ldots, n_i. \]  

(3)

Compared to Eq. (1), there is no restriction on the number of sub-populations that needs to be specified before model estimation. The actual number of sub-populations will be automatically determined from the available data jointly with the model estimation.

2.2. Model estimation

Estimation of the proposed model in Eq. (3) is mathematically nontrivial. Specifically, given the performance degradation data of \( N \) individuals, denoted as \( D = \{y_{ij}\}_{i=1}^{N} \), and a collection of unknown functions and parameters, denoted as \( \Theta = \{m_k(\cdot), \sigma_k^2\}_{k=1}^{\infty} \), the likelihood function \( L(\Theta | D) \) can be explicitly written as

\[ L(\Theta | D) = \prod_{i=1}^{N} \prod_{j=1}^{n_i} \left( \sum_{k=1}^{\infty} p_k f_k(y_{ij} | m_k(t_{ij}), \sigma_k^2) \right). \]  

(4)

Since the proposed model is a non-parametric Bayesian hierarchical model, a convenient Bayesian estimation approach needs to be developed. Under the Bayesian framework, the objective is to obtain the joint posterior density \( \pi(\Theta | D) \). Denoting \( \pi(\Theta) \) as the joint prior density function, the analytical derivation of \( \pi(\Theta | D) = L(\Theta | D) \cdot \pi(\Theta) / \pi(D) \) involves evaluating \( \pi(D) = \int \pi(\Theta, D) d\Theta \), which is mathematically intractable due to the high-dimensional integration. Markov chain Monte Carlo (MCMC) sampling algorithms need to be developed to draw exact samples for the joint posterior density \( \pi(\Theta | D) \). However, due to the complex form of \( L(\Theta | D) \) in Eq. (4), several estimation challenges need to be addressed: (1) All unknown parameters (e.g., \( \sigma_k^2 \)) and functions \( m_k(\cdot) \) are highly dependent. Such high dependency will result in failed convergence of conventional MCMC (Robert et al., 1997); (2) there are infinitely many parameters and functions to be estimated in \( \Theta \), which makes model estimation computationally formidable; (3) parametric form of \( m_k(\cdot) \) needs to be specified by considering both model flexibility and computational convenience.

To address the aforementioned model estimation challenges, a data augmentation technique is first adopted through augmenting auxiliary variables to simplify the complex dependency structure in \( L(\Theta | D) \). It has been successfully considered in both deterministic optimization for maximizing a likelihood function (Watanabe et al., 2003) and stochastic algorithm for sampling a posterior density (Tanner et al., 1987). Specifically, considering auxiliary variables \( Z = \{z_i\}_{i=1}^{N} \), where \( z_i \) is a categorical variable taking positive integers to represent the sub-population membership of individual \( i \), \( z_i = k \) indicates that performance degradation of individual \( i \) belongs to sub-population \( k \). It is noted that, in the proposed model, both the sub-population memberships \( z_i \) and the total number of sub-populations are unknown quantities. They will be simultaneously estimated together with the model parameters during the course of model estimation. Based on the data augmentation, the augmented joint likelihood function \( L(\Theta | Z, D) \) can be written as

\[ L(\Theta | Z, D) = \prod_{i=1}^{N} \prod_{k=1}^{\infty} \left( \prod_{j=1}^{n_i} f_k(y_{ij} | \phi_k(t_{ij}), \beta_k, \sigma_k^2) \right) I(z_i = k), \]

(5)

where \( I(\cdot) \) is an indicator function. With data augmentation, unlike Eq. (4), where all unknown functions and parameters are dependent, \( (m_k(\cdot), \sigma_k^2) \) become conditionally independent of \((m_k(\cdot), \sigma_k^2)\), \( \forall k \neq k' \), and the dependency structure is greatly simplified.

For sub-population \( k \), \( m_k(t) \) represents the mean performance degradation. To characterize the potential nonlinear-evolving degradation patterns while maintaining the Bayesian computation convenience, a linear additive functional form is specified; i.e., \( m_k(t) \approx \beta_k^T \phi_k(t) \). \( \phi_k(t) \) is a vector of basis functions, such as polynomial basis and splines, and \( \beta_k \) is the corresponding vector of coefficients. Based on such a specification, \( m_k(t) \) becomes a nonlinear function of \( t \) with adequate modeling flexibility in capturing various sub-population specific degradation patterns. On the other hand, \( m_k(t) \) is also a linear function of unknown coefficients \( \beta_k \), and the conjugate prior distributions are available to facilitate the Bayesian estimation. Specifically, considering the prior distribution for \( \beta_k \) as the multivariate normal distribution (i.e., \( \beta_k \sim \text{MVN}(\mu_k, \Sigma_k) \)), the corresponding full posterior conditional distribution can be expressed as

\[ \beta_k | \sigma_k^2, Z, D \sim \text{MVN}(\mu_k', \Sigma_k'), \]

(6)

where \( \mu_k' = \Sigma_k' (\Sigma_k^{-1} \mu_k + \Sigma_k^{-1} \mu_k) \) and \( \Sigma_k' = \Sigma_k^{-1} - \Sigma_k^{-1} \Phi_k(t) \Phi_k(t) \Sigma_k^{-1} \). \( \Phi_k(t) \) is a matrix with its \( j \)th row defined as \( \phi_j(t) \). \( Y_k \) is a matrix with its \( j \)th column represented as \( y_j, i \in S_k \), where index set \( S_k \) is defined as \( S_k = \{ i : z_i = k, \forall i = 1, \ldots, N \} \). Similarly, considering the prior distribution for \( \sigma_k^2 \) as the inverse gamma distribution (i.e., \( \sigma_k^2 \sim \text{IG}(a_k', b_k') \)), the corresponding full conditional posterior distribution can be expressed as

\[ \sigma_k^2 | \beta_k, Z, D \sim \text{IG}(a_k' + |S_k| + a_k, b_k' + \frac{1}{2} (Y_k - \Phi_k(t) \beta_k)^T (Y_k - \Phi_k(t) \beta_k) + b_k). \]

(7)

where \( a_k' = \frac{1}{2} |S_k| + a_k \), \( b_k' = \frac{1}{2} (Y_k - \Phi_k(t) \beta_k)^T (Y_k - \Phi_k(t) \beta_k) + b_k \) and \( | \cdot | \) denotes the cardinality of set. As shown in Eqs. (6) and (7), it is convenient to draw samples for both \( \beta_k | \sigma_k^2, Z, D \) and \( \sigma_k^2 | \beta_k, Z, D \) since sampling routines for such common distributions are readily available for most of the programming and computing environments.

Another estimation challenge to be solved is that for the augmented variables \( Z = \{z_i\}_{i=1}^{N} \). Each \( z_i \) could take any positive integer; i.e., \( z_i = 1, 2, \ldots \). In addition, an infinite number of parameters (i.e., \( \beta_k, \sigma_k^2 \)) are required to be estimated. To address such infinite-dimensionality issue, a slice sampling technique (Li et al., 2017; Walker, 2007) is considered. Specifically, introducing uniformly distributed random variables \( u_i \) (i.e., \( u_i | z_i = k \sim \text{Unif}(0, p_k) \)), the full conditional posterior for \( z_i \) becomes

\[ \pi(z_i = k | u_i, \Theta, D) \propto I(k \in B(u_i)) \prod_{j=1}^{n_i} f_k(y_{ij} | \phi_k(t_{ij}), \beta_k, \sigma_k^2), \]

(8)
where \( B(u_i) = \{k : p_k > u_i\} \). It is noted that \( p_k \) is the proportion of sub-population \( k \) satisfying \( \sum_{k=1}^{\infty} p_k = 1 \), which can be generated using the stick-breaking procedure as in the work of Sethuraman (1994); \( p_k = v_k \prod_{k=1}^{k-1}(1-v_k), k > 2, p_1 = v_1 \) and \( v_k \sim \text{Beta}(1, \eta) \). Based on the stick-breaking procedure, it can be shown that \( p_k < 1 - \sum_{k=1}^{k-1} p_k, \forall k \). Thus, when \( 1 - \sum_{k=1}^{K^*} p_k < \min\{u_i\}_{i=1}^{N} \), there is \( k \notin B(u_i), \forall k > K^* \). Then, the cardinality \( |B(u_i)| \leq k^* \) is a finite integer and \( z_i \) in Eq. (5) can only take \( |B(u_i)| \) number of distinct positive integers. Similarly, when \( 1 - \sum_{k=1}^{K^*} p_k < \min\{u_i\}_{i=1}^{N} \), there is \( k \notin \bigcup_{i} B(u_i), \forall k > K^* \). This implies that even though there is an infinite number of sub-populations in the formulation, in actual model estimation only \( K^* \) number of sub-populations need to be sampled.

Based on \( Z = \{z_i\}_{i=1}^{N} \) sampled in Eq. (8), the stick-breaking procedure of generating \( p_k|Z \) can be achieved based on the updated \( v_k|Z \), which is given by

\[
v_k|Z \sim \text{Beta}\left(1 + \sum_{i=1}^{N} I(z_i = k), N + \eta - \sum_{k=1}^{L} \sum_{i=1}^{N} I(z_i = l)\right). \tag{9}
\]

where \( \eta \) is the aforementioned positive diffusion value of DP.

In summary, the proposed estimation procedure for sub-population-level heterogeneous degradation modeling is shown in Algorithm 1.

**Algorithm 1** Sampling algorithm for proposed model.

Initialization: \( K^{(0)} = 1, z_i^{(0)} = 1, \forall i = 1, \ldots, N, (\beta_k^{(0)}, \sigma_k^{(2)(0)}) \sim \pi(\beta_k, \sigma_k^2), p_k^{(0)} = v_1^{(0)} \sim \text{Beta}(N + 1, \eta) \)

**procedure** DrawSamples

for \( \tau \leftarrow 1, \tau_{\text{max}} \)

\[u_1^{(\tau)}|z_1^{(\tau-1)} = k \sim \text{Unif}(0, p_k^{(\tau-1)})\]

obtain \( K^* \) from \( K^* = \min\{K : \sum_{k=1}^{K} p_k > 1 \}

\[-\min\{u_i\}_{i=1}^{N}\}, \text{set } K^{(\tau)} = K^{(\tau-1)}\]

if \( K^* > K^{(\tau-1)} \) then

\[K^{(\tau)} = K^*\]

generate \( v_k^{(\tau)} \sim \text{Beta}(N + 1, \eta) \) and compute

\[p_k^{(\tau)} = \prod_{k=1}^{K-1} (1 - v_k^{(\tau)}) v_k^{(\tau)}, \forall k > K^{(\tau-1)}\]

generate \( (\beta_k^{(\tau)}, \sigma_k^{(2)(\tau)}) \sim \pi(\beta_k, \sigma_k^2) \)

end if

update \( z_i^{(\tau)} = k|u_1^{(\tau)} \) by Eq. (8)

update \( (\beta_k^{(\tau)}, \sigma_k^{(2)(\tau)})|Z, D \) by Eqs. (6) and (7)

update \( v_k^{(\tau)}|Z \) by Eq. (9) and compute

\[p_1^{(\tau)} = v_1^{(\tau)}, p_k^{(\tau)} = \prod_{k=1}^{K-1} (1 - v_k^{(\tau)}) v_k^{(\tau)}, \forall k > 2\]

end for

end procedure

### 2.3. Individual-level heterogeneity modeling

In the previous section, based on the proposed model estimation algorithm, \( K \) number of sub-populations is identified jointly with the estimated model parameters \( \hat{\Theta} = (\hat{\beta}_k, \hat{p}_k, \hat{\sigma}_k^2, k = 1, \ldots, K) \). For the overall heterogeneous population, sub-population-level degradation heterogeneity is explicitly quantified by \( \hat{m}_k(t) \) with \( \hat{m}_k(t) = \hat{\beta}_k^T \hat{\phi}_k(t) \). The unknown sub-population membership for any individual \( i \) can be calculated as

\[
\hat{z}_i = \max_k \left\{ \lambda_k \hat{m}_k(t) | D \right\} = \max_k \frac{\hat{\beta}_k^T \hat{\phi}_k(t)}{\sum_k \hat{\beta}_k^T \hat{\phi}_k(t)}.
\]
satisfying \( \int \psi^2_q(t)dt = 1 \) and \( \int \psi_q(t)\psi^q(t)dt = 0, \forall q \neq q' \). Equation (12) indicates that the variability of \( \Delta X(t) \) can be projected onto several orthonormal eigenfunctions \( \psi_q(t) \), where \( \psi_1(t) \) carries the largest variability, \( \psi_2(t) \) carries the second-largest variability, and so forth. As \( q \) increases, the corresponding \( \psi_q(t) \) carries less variability and is less interpretable. Thus, the first \( Q \) number of eigenfunctions is often considered and the covariance function can be approximated as: \( \text{cov}(\Delta X(t), \Delta X(s)) \approx \sum_{q=1}^{Q} \lambda_q \psi_q(t)\psi_q(s) \).

Based on a finite number of eigenfunctions and according to the Karhunen–Loève theorem (Ramsay et al., 2005), the individual degradation heterogeneity \( \Delta x_i(t) \) of individual \( i \) can be approximately written as

\[
\Delta x_i(t) \approx \sum_{q=1}^{Q} \xi_{iq} \psi_q(t),
\]

where \( \xi_{iq} \) is the functional principal component score for the \( q \)th eigenfunction of individual \( i \), which is normally distributed with mean 0 and variance \( \lambda_q \). It controls the magnitude of \( \psi_q(t) \) contributing to \( \xi_{iq} \). Equation (13) is interpreted as follows. The individual degradation heterogeneity can be expressed as a weighted combination of multiple eigenfunctions while each eigenfunction \( \psi_q(t) \) characterizes a specific temporal variation signature of performance degradation.

To estimate the aforementioned eigenvalues, eigenfunctions, and scores of individuals, the performance degradation data \( D = \{y_{ij}, t_{ij}, \forall i=1, \ldots, N, j=1, \ldots, n_i\} \) of the heterogeneous population is first partitioned into sub-population degradation data \( D_{k,s} \), i.e., \( D_k = \{y_{ij}, \forall j = k, j=1, \ldots, n_i\} \), according to Eq. (10). For sub-population \( k \), a set of sub-population specific eigenvalues and eigenfunction pairs \( \{\lambda^{(k)}_q, \psi^{(k)}_q(t)\} \) can be estimated by maximizing the following likelihood function \( L(\lambda^{(k)}_q, \psi^{(k)}_q(t))\) as

\[
L(\lambda^{(k)}_q, \psi^{(k)}_q(t)) \propto \prod_{s=k} \frac{1}{\sqrt{\det(\Sigma^{(k)}_{\Delta t})}} \exp \left[ -\frac{1}{2} (y(t_i) - m_k(t_i))^{\top} \Sigma^{-1}_{\Delta t} (y(t_i) - m_k(t_i)) \right],
\]

where \( y(t_i) = [y_{i1}, \ldots, y_{in_i}]^{\top} \), \( m_k(t_i) = [m_{k1}(t_i), \ldots, m_{kn_i}(t_i)]^{\top} \) and \( \Sigma^{(k)}_{\Delta t} = \text{var}(y(t_i) - m_k(t_i)) \). The restricted Maximum likelihood estimation method proposed by Peng et al. (2009) can be adopted to estimate \( \{\lambda^{(k)}_q, \psi^{(k)}_q(t)\} \) as \( \{\hat{\lambda}^{(k)}_q, \hat{\psi}^{(k)}_q(t)\} \). The scores \( \hat{\xi}_{iq} \) of functional principal components can be further calculated (Yao et al., 2005) as: \( \hat{\xi}_{iq} = \sum_{j=1}^{n_i} (y_{ij} - m_k(t_{ij})) \hat{\psi}^{(k)}_q(t_{ij}) (t_{ij} - t_{i,j-1}) \) and \( t_0 = 0, \forall i = 1, \ldots, N, q = 1, \ldots, Q \).

### 3. Case study

#### 3.1. Numerical case study

To demonstrate the effectiveness of the proposed Bayesian sub-population-level heterogeneity quantification method, a heterogeneous population with two sub-populations is assumed. Third-order polynomial basis functions (i.e., \( \phi(t) = [1, t, t^2, t^3]^{\top} \)), are considered for illustration purposes. For simplicity, sub-population \( k \) is written as “sub-\( k \” in short. Ground-truth values for sub-population \( k \)'s \((k = 1, 2)\) parameters, namely the mixing proportion \( \pi_k \), a vector of basis coefficients \( \beta_k = [\beta_{1k}, \beta_{2k}, \beta_{3k}]^{\top} \) and the within-sub-population variability \( \sigma_k^2 \), are pre-assigned and summarized in Table 1. A total of 1000 samples are simulated for this study.

Figure 2 shows trace plots and histogram of the results of the estimated number of sub-populations. The estimated number of sub-populations over iterations stabilizes at two in Fig. 2a and the histogram highly concentrates at two sub-populations in Fig. 2b. They both indicate that the identified number of sub-population is two, which is consistent with the ground-truth specification of two sub-populations. Deviance values are further displayed in Fig. 2c to evaluate the goodness-of-fit of the identified model over iterations and to monitor the convergence (Green et al., 2001) of the sampling algorithm. The algorithm converges fast within 20 iterations since conjugate priors greatly facilitate the estimation convenience and performance. The corresponding point (e.g., posterior mean) and interval (e.g., 95% credible intervals) estimation results are summarized in Table 1. Maximum likelihood estimation (MLE) results assuming a known ground-truth number of sub-populations and known sub-population membership for each observation are also shown in Table 1. The posterior means of all parameters are almost identical to the MLE results, since Bayesian estimation is dominated by the available data rather than prior knowledge. Both estimation results are also close to the pre-assigned ground-truth values. Some discrepancies can be seen for sub-population 2 due to its large within-sub-population variability (i.e., \( \sigma_2^2 = 100 \)), and a finite sample size of data. Bayesian 95% credible intervals successfully quantify

<table>
<thead>
<tr>
<th>Parameters</th>
<th>True value</th>
<th>Maximum likelihood estimation</th>
<th>Posterior mean</th>
<th>95% Credible interval</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \pi_1 )</td>
<td>0.4</td>
<td>0.39</td>
<td>0.39</td>
<td>[0.36, 0.42]</td>
</tr>
<tr>
<td>( \pi_2 )</td>
<td>0.6</td>
<td>0.61</td>
<td>0.61</td>
<td>[0.58, 0.64]</td>
</tr>
<tr>
<td>( \beta_{11} )</td>
<td>1</td>
<td>1.09</td>
<td>1.09</td>
<td>[1.02, 1.16]</td>
</tr>
<tr>
<td>( \beta_{12} )</td>
<td>1</td>
<td>0.33</td>
<td>0.33</td>
<td>[0.29, 0.37]</td>
</tr>
<tr>
<td>( \beta_{21} )</td>
<td>0.2</td>
<td>0.10</td>
<td>0.10</td>
<td>[0.06, 0.14]</td>
</tr>
<tr>
<td>( \beta_{22} )</td>
<td>0.4</td>
<td>0.96</td>
<td>0.96</td>
<td>[0.90, 1.02]</td>
</tr>
<tr>
<td>( \beta_{31} )</td>
<td>0.3</td>
<td>0.33</td>
<td>0.33</td>
<td>[0.29, 0.37]</td>
</tr>
<tr>
<td>( \beta_{32} )</td>
<td>0.2</td>
<td>0.07</td>
<td>0.07</td>
<td>[0.04, 0.10]</td>
</tr>
<tr>
<td>( \sigma_1^2 )</td>
<td>25</td>
<td>24.70</td>
<td>24.70</td>
<td>[23.70, 25.89]</td>
</tr>
<tr>
<td>( \sigma_2^2 )</td>
<td>100</td>
<td>101.00</td>
<td>100.98</td>
<td>[97.45, 104.57]</td>
</tr>
</tbody>
</table>
such estimation uncertainty and fully cover all ground-truth values. The numerical case study can also help investigate an important issue regarding the miss-clustering of the sub-populations memberships. Unlike the real case study, where the ground-truth values of sub-population memberships are unavailable, the ground-truth membership values of simulated data are known and it is possible to quantify the miss-clustering error as the miss-classification error. By comparing the predetermined ground truth values $z_i$ and the estimated values $\hat{z}_i$ in Eq. (10), the proposed method achieves 100% accuracy of sub-population memberships.

To further investigate the influence of hyper-parameter $\eta$ on the estimation performance, Fig. 3 shows both trace plots and deviance monitoring results under scenarios of $\eta = 2$, $\eta = 10$ and $\eta = 50$, respectively. They both correctly identify the ground-truth number of sub-populations, and the numbers of iterations required to converge are also similar. Since the influence of hyper-parameter $\eta$ is negligible, $\eta = 2$ is chosen to show the estimation results of Table 1 for illustration purposes.

### 3.2. Real case study

#### 3.2.1. Data description

To demonstrate the performance of the proposed heterogeneity quantification framework, the biennial longitudinal survey data of the Health and Retirement Study (HRS) is considered. The HRS data contains a wealth of information on disability and health conditions, socio-demographics, and health service utilization of a representative sample of community-dwelling older adults in the United States (Institute for Social Research, 2016). In this article, a subset of cognitive degradation data of 2457 individuals with complete information is investigated, since cognitive degradation is a hallmark of aging-related functional impairment and development of disability. The measurement metric of cognition of the HRS data was developed based on derivatives of the MMSE score (Folstein et al., 1975). A telephone questionnaire containing 35 questions with binary outcomes (i.e., 1 indicating affirmative response of good health and 0 indicating otherwise) was given to elderly individuals in eight waves from 1998 to 2012 on a biennial basis. Such questions evaluate the multi-dimensional cognitive capabilities of the older adults, such as memory, information remembrance, comparative reasoning, judgment and behavioral consistency. The total score (ranging from 0–35) of the questionnaire is calculated as a cognition index to assess the overall cognitive performance (Ofstedal et al., 2005). The proposed work aims to investigate both the sub-population-level and the individual-level heterogeneity of cognitive degradation.

#### 3.2.2. Sub-population-level cognitive degradation heterogeneity

To model the sub-population-level heterogeneity of cognitive degradation, the proposed non-parametric modeling approach is implemented to obviate the requirement of pre-specifying the number of sub-populations. Three-order basis functions are considered to capture the sub-population mean degradation trajectory. According to existing literature (Nagin, 2005; Nagin et al., 2010) and visual inspection of real data, three-order basis functions are reasonable choices in representing the overall trend of mean degradation trajectory. Modeling degradation heterogeneity at the individual level may require higher-order basis functions to capture detailed evolving patterns among individuals, which will be addressed by FPCA in Section 3.2.3. Compared to individual degradation trajectory, the mean degradation trajectory is at a more aggregate level to explore the similar evolving trends among individuals, thus requiring less modeling fidelity. Figure 4 shows the estimation results of the number

![Figure 3. Comparison of results with different $\eta$.](image-url)
of sub-populations $K$. The algorithm converges approximately at 4500 iterations, and the identified number of sub-populations is 3. Figure 5 further displays the mean degradation trajectories (i.e., $\hat{\beta}_k^T \phi_k(t)$) and their corresponding variability bounds (i.e., $\hat{\beta}_k(t) \pm 2\hat{\sigma}_k$). The mean trajectory of each sub-population describes the average degradation of cognitive condition and the corresponding variability bounds describe the within-sub-population variability. As shown in Fig. 5, the average cognition index values for sub-populations 1, 2 and 3 at the baseline wave (i.e., wave 1) are 27.5, 24.5 and 20.5, while their corresponding average absolute decrements after eight waves are 1.5, 2 and 5, respectively. Thus, the cognitive degradation of the overall heterogeneous aging population can be characterized by three sub-populations with low, median and high levels of cognitive impairment in sub-1, 2 and 3, respectively.

The proposed work is also compared with some of the existing parametric modeling approaches (Jones et al., 2007; Nagin et al., 2010) in investigating the sub-population-level heterogeneity. Parametric models require a two-step model estimation and selection procedure. In the model estimation step, candidate models which pre-specify a fixed number of sub-populations are developed and estimated. Then, in the model selection step, the final model is selected among the candidate models based on an information criterion, such as AIC and BIC. Table 2 summarizes the comparative results. All three methods indicate the existence of sub-population-level heterogeneity in cognitive degradation. Different methods select different numbers of sub-populations, or equivalently, admit different model complexities in the context of heterogeneity modeling. The BIC tends to select a simpler model with a smaller number of sub-populations since more penalty is imposed to control the model complexity. In contrast, the AIC tends to select a more complex model with a larger number of sub-populations since less penalty term for controlling the model complexity is involved. The model complexity (or the number of sub-populations) of the proposed method is automatically determined by the data. There is no need to pre-specify the number of sub-populations or pre-design the penalty structure in controlling the model complexity. Moreover, unlike two-step modeling approaches, model estimation and selection can be accomplished simultaneously in a single step in the proposed work.

### 3.2.3. Individual-level cognitive degradation heterogeneity

In the previous section, a non-parametric modeling approach is proposed and applied to characterize the sub-population heterogeneity of cognitive degradation. The overall heterogeneous aging population is decomposed into three groups of individuals with different average levels (i.e., low, median and high) of cognitive performances and evolving patterns. As shown in Fig. 5, for each of the identified sub-population, there is a significant amount of within-sub-population variability.

<table>
<thead>
<tr>
<th>Variables</th>
<th>AIC based</th>
<th>BIC based</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\nu_k$</td>
<td>sub-1: 0.226</td>
<td>sub-1: 0.198, 0.256</td>
</tr>
<tr>
<td></td>
<td>sub-2: 0.505</td>
<td>sub-2: 0.480, 0.531</td>
</tr>
<tr>
<td></td>
<td>sub-3: 0.261</td>
<td>sub-3: 0.228, 0.265</td>
</tr>
<tr>
<td>$\beta_{0k}$</td>
<td>sub-1: 28.060</td>
<td>sub-1: 27.463, 28.701</td>
</tr>
<tr>
<td></td>
<td>sub-2: 24.803</td>
<td>sub-2: 24.365, 25.262</td>
</tr>
<tr>
<td></td>
<td>sub-3: 21.267</td>
<td>sub-3: 20.479, 22.045</td>
</tr>
<tr>
<td>$\beta_{1k}$</td>
<td>sub-1: -0.561</td>
<td>sub-1: -1.087, -0.044</td>
</tr>
<tr>
<td></td>
<td>sub-2: -0.536</td>
<td>sub-2: -0.912, -0.360</td>
</tr>
<tr>
<td></td>
<td>sub-3: -0.952</td>
<td>sub-3: -1.639, -0.242</td>
</tr>
<tr>
<td>$\sigma^2_k$</td>
<td>sub-1: 7.458</td>
<td>sub-1: 7.082, 7.853</td>
</tr>
<tr>
<td></td>
<td>sub-2: 8.560</td>
<td>sub-2: 8.265, 8.847</td>
</tr>
<tr>
<td></td>
<td>sub-3: 15.711</td>
<td>sub-3: 15.016, 16.298</td>
</tr>
</tbody>
</table>

Note: Sub-1, sub-2, sub-3 and sub-4 are abbreviations for sub-population 1, 2, 3, and 4, respectively.
that has not been further explained and quantified. Such within-sub-population variability is caused by individual level degradation heterogeneity nested within each sub-population. Based on the modeling outputs from the previous section, FPCA is employed in this section to investigate the individual-level heterogeneity by treating each individual degradation trajectory as a functional data object.

In the context of cognitive degradation modeling, FPCA essentially projects the within-sub-population variability into a set of orthonormal eigenfunctions. Each individual degradation trajectory can be represented by a weighted combination of these eigenfunctions. The individual specific weights, defined as FPCA scores, can be utilized to quantify the individual degradation heterogeneity. Considering the most parsimonious number of eigenfunctions which gives desirable interpretation and retains a reasonable portion of variability, two eigenfunctions are selected for each sub-population. Based on the approximate leave-one-curve-out cross-validation scores (Peng et al., 2009), four-order basis functions of splines are selected for eigenfunctions within sub-populations 1 and 3, and six-order for sub-population 2 to avoid over-fitting the data. Figure 6 summarizes the sub-population specific eigenfunctions and visualizes their meaningful interpretations through comparison with the mean degradation trajectories (in solid curves). For simplicity, the eigenfunction of the first principal component for sub-population 1 is termed as EPC1-1 in short and similar abbreviations are applied to other eigenfunctions.

As shown in Fig. 6a, a positive/negative score of EPC1-1 indicates an above-average/below-average cognitive degradation over the entire time period. Its degradation rate is similar to the average degradation of sub-population 1. As shown in Fig. 6b, a positive score of EPC1-2 represents the above-average degradation performance value at the beginning but below-average performance value in the end. In addition, its
degradation first accelerates and then decelerates compared to the mean degradation rate of sub-population 1. As opposed to EPC1-1, for EPC1-2 in Fig. 6c, a positive/negative score indicates a below-average/above-average cognitive degradation. For EPC2-2, after above-average degradation transits into below-average degradation between waves 4 and 5, the cognition index value will first decrease then become stabilized. However, for EPC3-2 in Fig. 6f, the cognition index value will continue to decrease monotonically, indicating a higher level of cognitive impairment and deterioration of individuals from sub-population 3. In addition, the transition times for EPC2-2 and EPC3-2 are slightly delayed than that of EPC1-2.

Based on FPCA, the individual-level degradation heterogeneity can be directly quantified by the lower-dimensional scores associated with different eigenfunctions. Figure 7 visualizes the scores of EPC1-1 and EPC1-2 for all individuals belonging to sub-population 1. Individual cognitive degradation heterogeneity can be concisely summarized and visualized in a two-dimensional plane. These concise scores will serve as informative and quantitative indices in helping LTC decision makers better allocate available nursing resources for individuals. Taking individuals A, B, C and D as four extreme cases, the positive score of EPC1-1, the negative score of EPC1-1, the positive score of EPC1-2, and the negative score of EPC1-1 are dominated, respectively. Based on the interpretation of eigenfunctions, the extreme scores indicate that individual A will have an above-average degradation, individual B will have a below-average degradation, individual C will first have an above-average then below-average degradation, and individual D will have a first below-average then above-average degradation.

In addition to the aforementioned concise and meaningful interpretations offered by FPCA, the modeling accuracy can be improved over the existing individual heterogeneity modeling approaches, such as growth-curve models. Specifically, we compare the modeling accuracy among FPCA and different growth-curve models with different individual-specific random parameters, such as the random intercept (RI), the random slope (RS), as well as the random intercept and slope (RIS). The overall mean square errors between predicted and observed degradation performances based on FPCA, RI, RS and RIS are 5.456, 7.273, 7.908 and 5.839, respectively. The superior modeling accuracy of the FPCA results is mainly because eigenfunctions capture more temporal information, such as higher-order degradation curvature. Figure 8 further visualizes the comparison results for the above representative individuals and demonstrates the satisfactory modeling accuracy of FPCA. The interpretation findings of FPCA can also be justified in Fig. 8.

To further explain the individual level heterogeneity of cognitive degradation within each sub-population, it will be desirable to investigate how FPCA scores are correlated with different observed individual covariates, such as health conditions, socio-demographics and healthcare financing status. Table 3 summarizes the significant and interpretable individual covariates for all three sub-populations based on the regression analysis. The maximum allowable significance level is set as 0.1. For sub-population 1, a higher value of EPC1-1 indicates a higher-than-average cognitive performance degradation over time. A risk/protective factor tends to decrease/increase the EPC1-1 value. All significant covariates for EPC1-1 have negative effects and can be treated as risk factors. They can be interpreted as: an individual, who is older at the initial interview, has a higher body mass index (i.e., higher level of obesity), covered under the Government Medicare insurance, and/or is not African American, Caucasian or Hispanic, will tend to experience a lower-than-average cognitive degradation over time.

![Figure 7](image_url) Scatter plot of EPC1-1 vs. EPC1-2 scores with extreme individuals identified.

![Figure 8](image_url) Comparison of observed, estimated and mean trajectories with growth-curve model results for: (a) high positive EPC1-1 score; (b) high negative EPC1-1 score; (c) high positive EPC1-2 score; (d) high negative EPC1-2 score.
Table 3. Post-analysis results of identified covariates explaining individual heterogeneity (standard errors, SE, in parentheses).

<table>
<thead>
<tr>
<th>Sub-population 1 covariates</th>
<th>EPC1-1</th>
<th>EPC1-2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age at initial interview in years</td>
<td>-0.061 (0.015)**</td>
<td>-0.023 (0.013).</td>
</tr>
<tr>
<td>Body mass index in KG/m²</td>
<td>-0.857 (0.289)**</td>
<td>-1.195 (-0.447)**</td>
</tr>
<tr>
<td>Government Medicare indicator (1 if insurance covered by Medicare, 0 otherwise)</td>
<td>-1.05E-06 (6.52E-07).</td>
<td>0.225 (0.128).</td>
</tr>
<tr>
<td>Other Race indicator (1 if not African American, Caucasian or Hispanic, 0 otherwise)</td>
<td>-1.05E-06 (6.52E-07).</td>
<td>0.225 (0.128).</td>
</tr>
<tr>
<td>Heart problem indicator (1 if ever had heart problem, 0 otherwise)</td>
<td>-0.061 (0.015)**</td>
<td>-0.023 (0.013).</td>
</tr>
<tr>
<td>Sub-population 2 covariates</td>
<td>EPC2-1</td>
<td>EPC2-2</td>
</tr>
<tr>
<td>Age at initial interview in years</td>
<td>0.034 (0.008)**</td>
<td>0.022 (0.007)**</td>
</tr>
<tr>
<td>Combined income of individual and spouse in USD</td>
<td>-1.46E-06 (8.46E-07).</td>
<td>-0.154 (0.071)*</td>
</tr>
<tr>
<td>Smoking indicator (1 if ever smoked, 0 otherwise)</td>
<td>-2.48E-06 (1.64E-06).</td>
<td>-2.24E-06 (1.09E-06).</td>
</tr>
<tr>
<td>Number of doctor visits in previous 2 years</td>
<td>-0.240 (0.097)*</td>
<td>0.0455 (0.190)*</td>
</tr>
<tr>
<td>Long-term care insurance indicator (1 if covered by LTC insurance, 0 otherwise)</td>
<td>-0.955 (0.143)**</td>
<td>-0.062 (0.299)*</td>
</tr>
<tr>
<td>Hispanic indicator (1 if Hispanic, 0 otherwise)</td>
<td>-0.955 (0.143)**</td>
<td>-0.062 (0.299)*</td>
</tr>
<tr>
<td>African American indicator (1 if African American, 0 otherwise)</td>
<td>-0.955 (0.143)**</td>
<td>-0.062 (0.299)*</td>
</tr>
<tr>
<td>Other Race indicator (1 if not African American, Caucasian or Hispanic, 0 otherwise)</td>
<td>-0.955 (0.143)**</td>
<td>-0.062 (0.299)*</td>
</tr>
<tr>
<td>Home health care indicator (1 if ever had health care at home in previous 2 years, 0 otherwise)</td>
<td>-0.955 (0.143)**</td>
<td>-0.062 (0.299)*</td>
</tr>
<tr>
<td>Sub-population 3 covariates</td>
<td>EPC3-1</td>
<td>EPC3-2</td>
</tr>
<tr>
<td>Age at initial interview in years</td>
<td>-0.034 (0.018).</td>
<td>0.079 (0.011)**</td>
</tr>
<tr>
<td>Total of all assets owned in USD</td>
<td>7.31E-07 (2.21E-07)**</td>
<td>-1.318E-07 (3.03E-07)**</td>
</tr>
<tr>
<td>Hospital stay indicator (1 if stayed overnight in hospital in previous 2 years, 0 otherwise)</td>
<td>-0.453 (0.218)*</td>
<td>-0.098 (0.413)*</td>
</tr>
<tr>
<td>Partial retiree indicator (1 if retired partially, 0 otherwise)</td>
<td>-0.791 (0.414).</td>
<td>-0.098 (0.413)*</td>
</tr>
<tr>
<td>African American indicator (1 if African American, 0 otherwise)</td>
<td>-1.82 (0.234)**</td>
<td>-0.098 (0.413)*</td>
</tr>
<tr>
<td>Other Race indicator (1 if not African American, Caucasian or Hispanic, 0 otherwise)</td>
<td>-0.098 (0.413)*</td>
<td>-0.098 (0.413)*</td>
</tr>
<tr>
<td>Number of living children</td>
<td>-0.06 (0.024)*</td>
<td>-0.06 (0.024)</td>
</tr>
</tbody>
</table>

Notes: p<0.1; * p<0.05; ** p<0.01; *** p<0.001.
95% Confidence Intervals for each parameter estimate are calculated by \( \hat{\beta}_j \pm 1.96 \times SE(\hat{\beta}_j) \), where \( \hat{\beta}_j \) is the respective estimated covariate coefficient.

4. Conclusion

A bi-level heterogeneity modeling and quantification framework is presented in this article to systematically and comprehensively characterize the heterogeneous functional performance degradation of older adults. A Bayesian non-parametric model is first presented to characterize heterogeneous degradation at the sub-population-level which bypasses the limitation of pre-specifying a known and fixed number of sub-populations in the conventional models. The resulting model is complex and makes conventional Bayesian estimation strategies challenging. An estimation algorithm is further developed to address the challenging issues of high dependency and infinite dimensionality in model estimation. Based on the proposed estimation algorithm, the number of sub-populations can be jointly identified with the parameters estimation in a one-step procedure. A numerical case study is provided to demonstrate the effectiveness of the proposed non-parametric model and compare it with parametric group-based modeling approaches. After quantifying the sub-population-level heterogeneity, FPCA is employed to extract temporal variation signatures of degradation profiles for individuals within each sub-population. The scores of FPCA quantify the individual heterogeneity and allow post-analysis in investigating the influence of possible individual covariates. To illustrate the proposed work, a real case study of heterogeneous cognitive degradation is provided, using the HRS data. The sub-population-level modeling identifies three distinct sub-populations with high, median and low levels of cognitive degradation. Within each sub-population, temporal variation signatures of individual heterogeneity are extracted and explained by several external factors, such as demographic and socioeconomic factors, health conditions and long-term care policies.
care supports. The proposed work successfully characterizes the heterogeneous functional performance degradation of the aging population. The modeling outputs will facilitate LTC policy-makers and administrators to better project heterogeneous LTC demand and utilization over time of older adults in a region of interest or in a LTC facility and to further develop more proactive, targeted and adaptive LTC preparedness and service delivery decision-making strategies at both LTC network and facility levels. In this article, the proposed Bayesian non-parametric model automatically identifies the number of sub-populations while the number of basis functions still needs to be pre-specified. One interesting future work is to further realize automatic and simultaneous identification of both the number of sub-populations and the number of basis functions required.

Funding
This work was supported in part by University of South Florida Research & Innovation Internal Awards Program under Grant No. 0114783.

References
Health and Retirement Study (HRS). (2016) *The Health and Retirement Study (HRS)*. Accessed at: https://hrs.isr.umich.edu/about